Analysis of the Value Added When Deploying a Model-Based Approach for the Validation and Verification of the Medical Database Software

Human Research Program
Exploration Medical Capability Element

HRP IWS
February 2023

Christopher Laing, Marina Parker, Tyler Duke, Mary Susan Kaetzer, Ronald Daiker, Michael Krihak, Kerry McGuire, Frank Cruz, Philip Augustine

Expanding the Boundaries of Space Medicine and Technology

Agenda

- **Introduction**
- **Informing Mission Planning via Analysis of Complex Tradespaces (IMPACT) / IMPACT Medical Database (MD)**
- **IMPACT MD's Validation and Verification (V&V) Efforts**
  - IMPACT MD V&V Plan
  - Validation Activity – Usability Testing
  - Findings and Outcomes
- **Conclusion**
- **Future Work**
- **Questions and Comments**

Introduction

- **Last year, the IMPACT Medical Database team presented on the innovative approaches that were implemented for IMPACT that include:**
  - Using Model-Based Systems Engineering (MBSE) to create a Concept of Operations (ConOps) model (vs. the traditional documentation approach).
  - Leveraging Human Factors' (HF) Human-Centered Design (HCD) approach to support software validation activities.

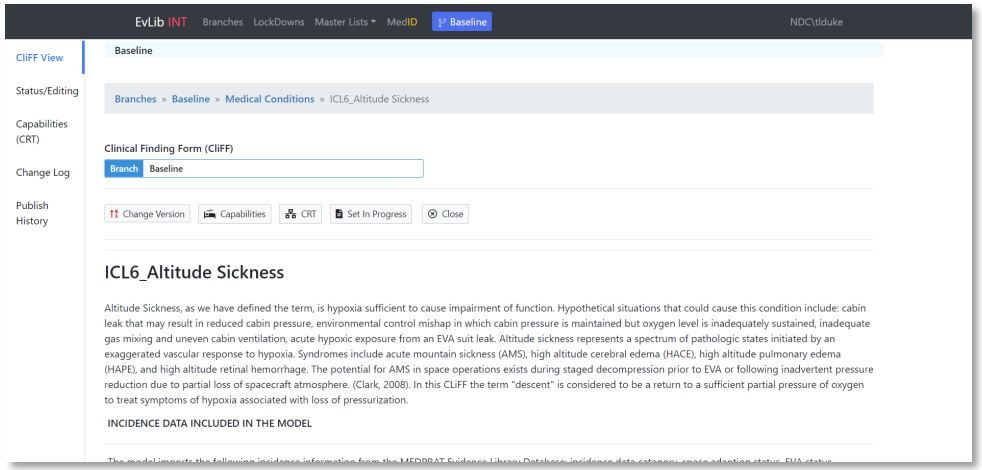
- **These approaches have allowed us to:**
  - Keep our users (active stakeholders) involved and at the forefront throughout the process.
  - Ensure an optimal product is being developed.
  - Reduce project costs, and coordination across project teams.

- **This presentation will focus on:**
  - A new approach to using MBSE products during usability testing.
  - The benefits of integrating MBSE and HF for software Verification & Validation efforts.

IMPACT Medical Database

- **IMPACT Medical Database (MD) is a software tool containing a database of medical evidence and engineering data.**
- **It provides information to the IMPACT tool suite necessary for users to make risk-informed decisions about medical systems for exploratory space missions.**
- **The MD tool supports an interface for clinicians to view, enter and modify data using the Evidence Library component**
- **The validation of this user interface is the focus of this presentation.**

Evidence Library



IMPACT MD V&V Plan

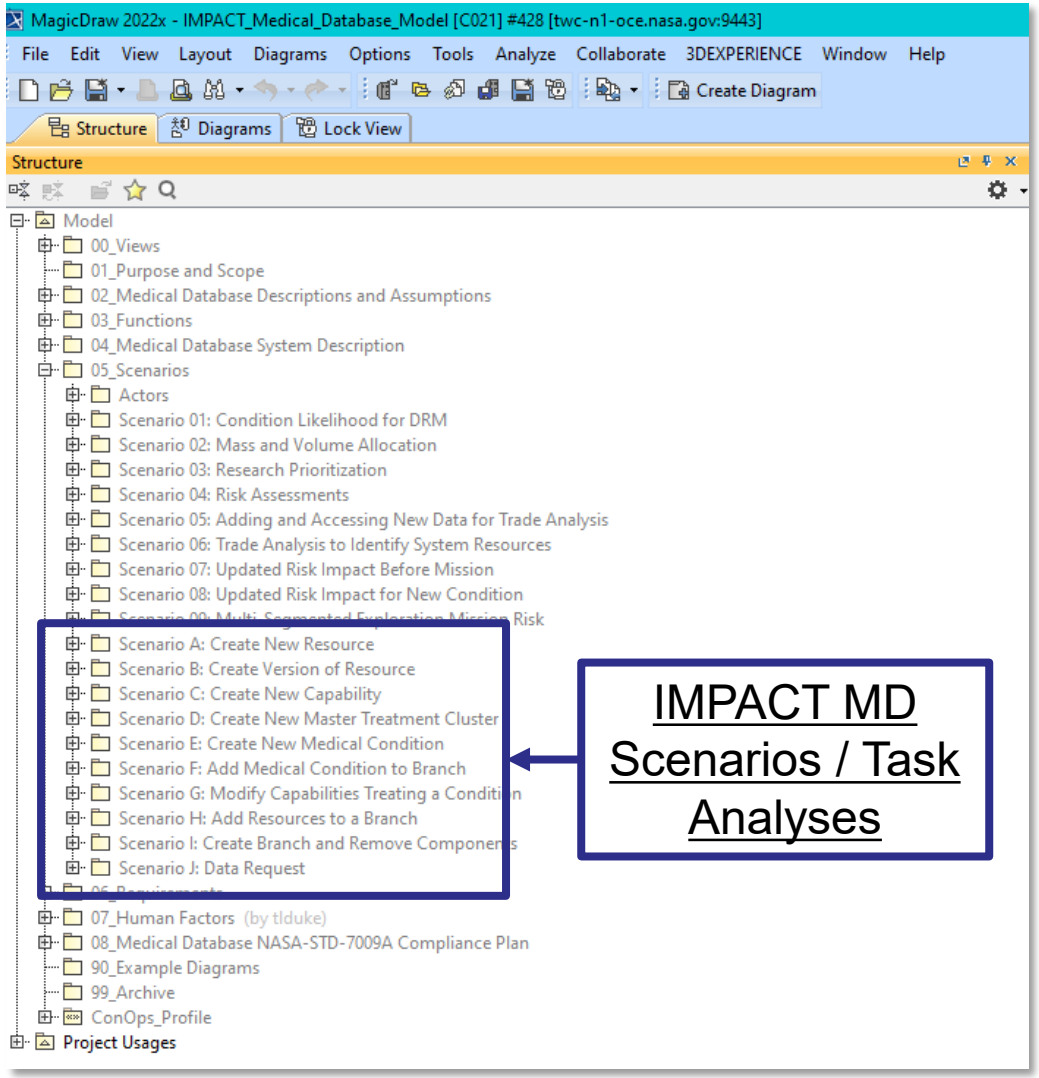
System certification for IMPACT (including the MD software) will be via the successful completion of the Transition to Operations Process as defined in Appendix D of NPR 8900.1B. This will require the IMPACT project team to produce a proposal with the following documentation:

1. A detailed description of the deliverable or product, its intended use or application, and a description of how the deliverable or product addresses a NASA-identified critical risk, medical, health, performance issues, or application.
2. Data demonstrating the efficacy, effectiveness, or utility of the deliverable or product.
3. Data demonstrating the operational validation of the deliverable or product.
4. An implementation plan of how the product or deliverable is to be used or applied (e.g., protocol, dosing regimen, scope of use).
5. An analysis of the mission resources (e.g., crew time, volume, power, etc.) necessary to implement the product or deliverable.
6. A summary of the developmental process and milestones the product or deliverable underwent.

IMPACT MD V&V Plan

- **IMPACT MD's software V&V plan was developed to utilize existing products (e.g., task analyses created in the MBSE tool, MagicDraw) and activities (e.g., usability testing).**
- **One effort to demonstrate the efficacy, effectiveness, and utility of the product, as well as the operational validation of the deliverable or product, involved utilizing the task analyses in MagicDraw to plan and perform usability tests with users.**
- **This is the effort that we are presenting today.**

MagicDraw



Validation Efforts – Usability Testing

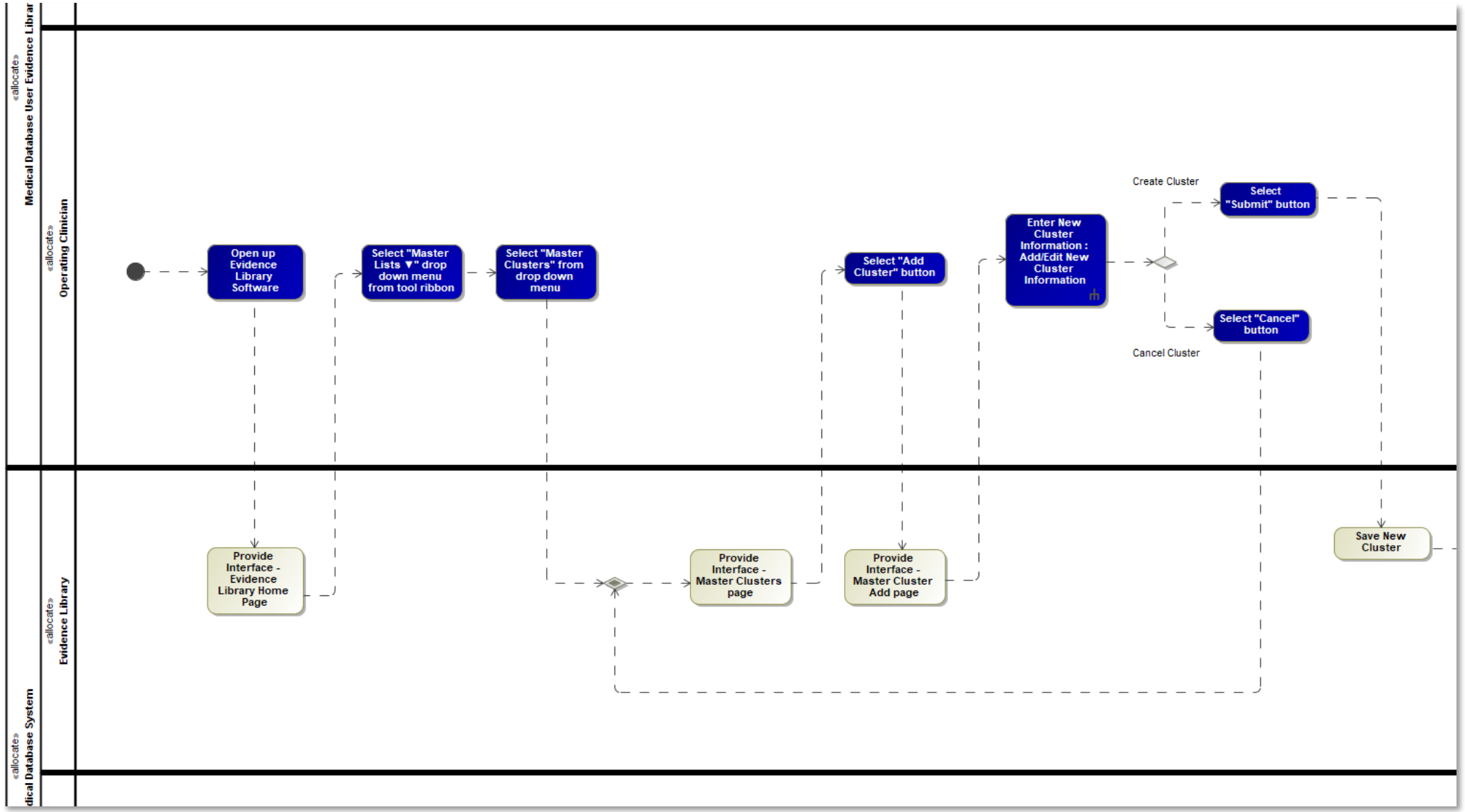
- **Usability testing is a method used to evaluate a product and involves representative users completing typical tasks while researchers observe and take notes.**
- **Some goals of usability testing include:**
  - Identifying usability problems.
  - Understanding users' perception of and satisfaction with the product.
  - Generating recommendations to improve the design and user satisfaction.
- ***Task analyses* are used to accomplish usability testing for IMPACT MD.**

Usability Testing Process

Validation Efforts – Usability Testing

- **The planning phase involves developing a test plan that includes:**
  - Scope, Purpose, and/or Objectives
  - Schedule, Location, and Equipment
  - Participants
  - Tasks/Scenarios
  - Metrics
- **This phase can be time-consuming, especially when developing representative tasks that users will perform. However, the human factors team leverages the task analyses that were developed using the MBSE tool MagicDraw.**

Validation Efforts – Usability Testing

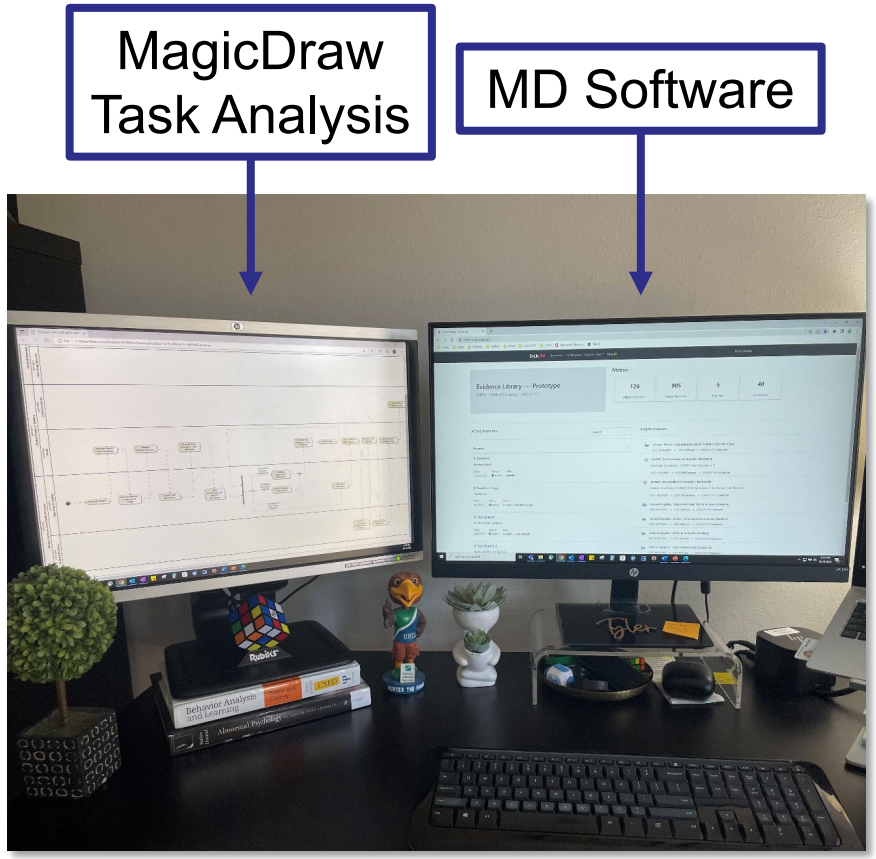


Validation Efforts – Usability Testing

- **During the next phase, users are recruited and scheduled to participate in the usability test. It is important to recruit:**
  - Representative users (i.e., users that are or have similar characteristics to expected end users) to ensure that the results obtained are effective.
  - A range of users (i.e., novice, intermediate, and advanced users) to ensure that different perspectives of the design and use of the software are considered.

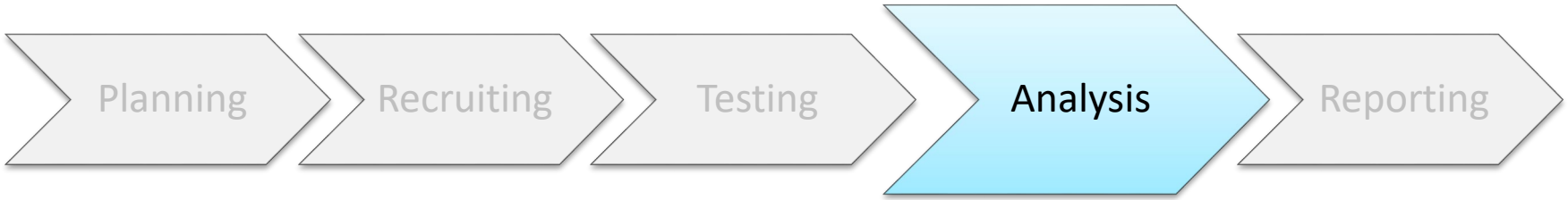
Validation Efforts – Usability Testing

- **The testing phase involves users performing each identified task/scenario in the test plan using the product and collecting data (e.g., test metrics and user feedback).**
- **A new approach to usability testing to support the validation of the IMPACT MD software (MedID and Evidence Library) was allowing participants to use the task analyses while performing tasks.**



Validation Efforts – Usability Testing

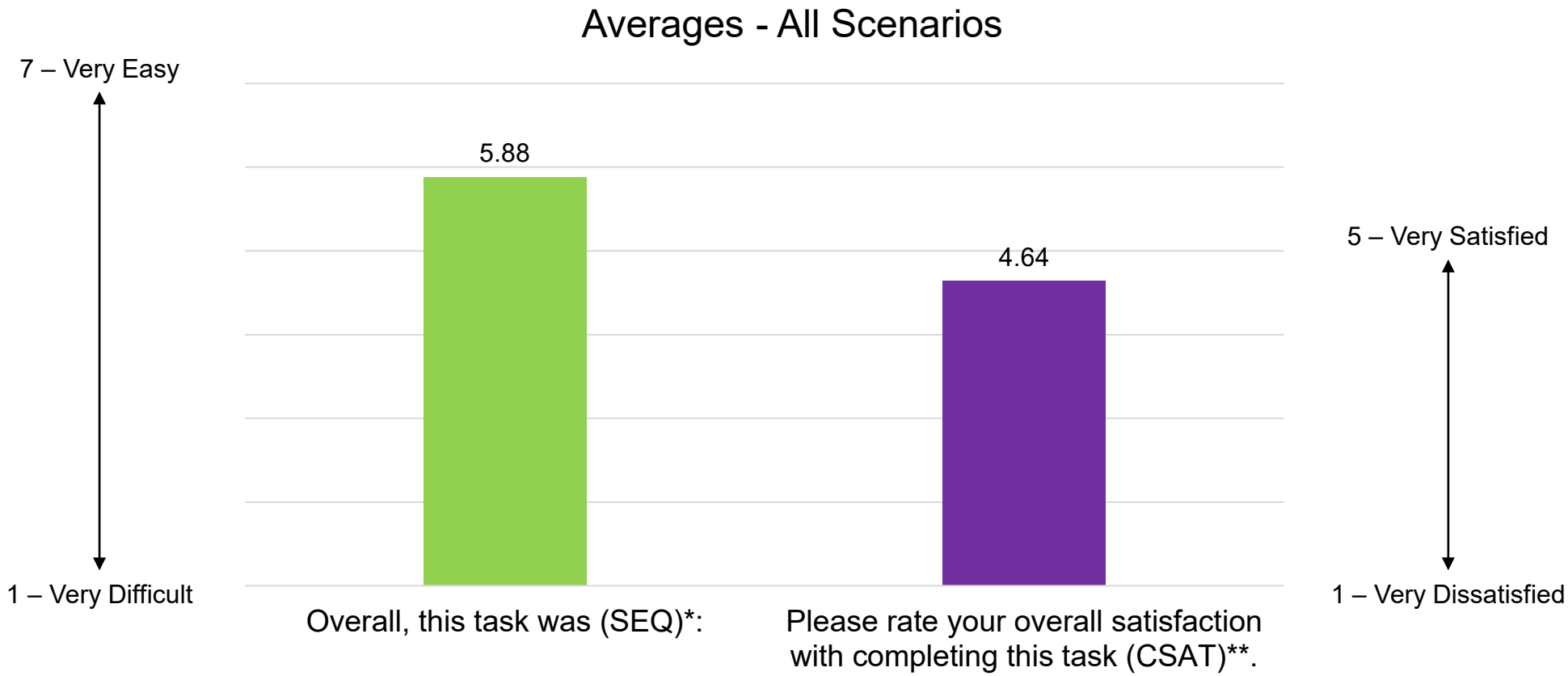
- **After testing with all participants, the data is then analyzed and synthesized.**
- **For IMPACT MD, this data includes:**
  - Average responses to post-task questions (ease of use and satisfaction)
  - Average responses to the System Usability Scale (SUS)
  - Task success/completion rate
  - Test observations and user feedback



Validation Efforts – Usability Testing

- **Finally, a report is generated that contains:**
  - Test objectives
  - Participant demographics
  - Methodology
  - Results
  - Recommendations
- **The human factors and developer teams discuss the recommendations and create a plan for implementation.**

Usability Test Findings – Post-Task Questionnaire



*A Single Ease Question (SEQ) score above 5.5 would be considered ***above average*** and anything below 5.5 is ***below average***.

A Customer Satisfaction Score (CSAT) score above 3 would be considered ***above average* and anything below 3 is ***below average***.

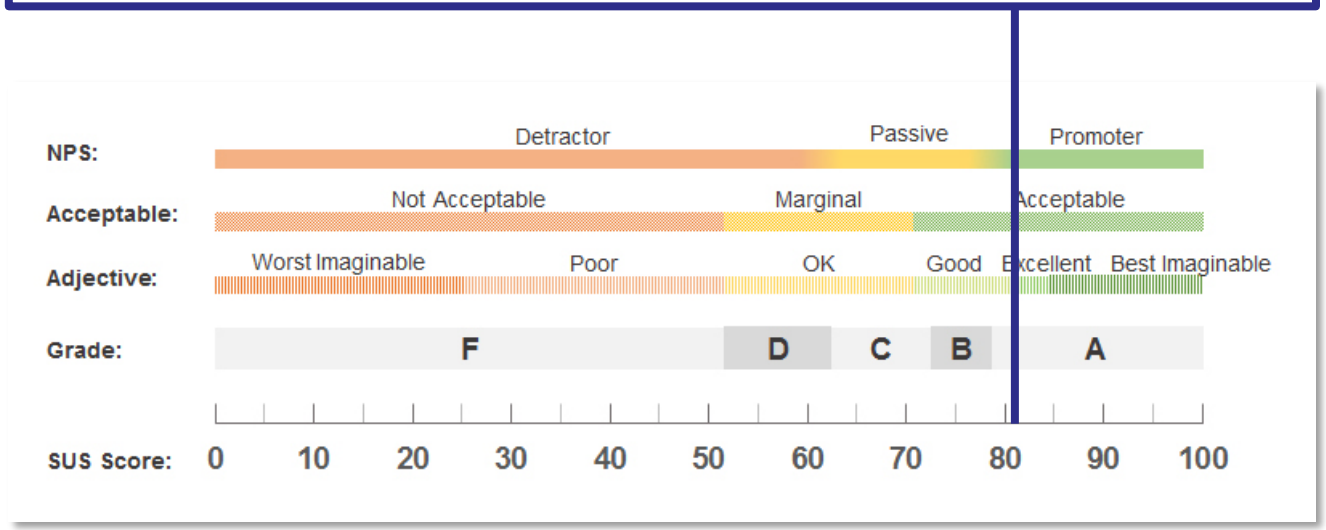
Usability Test Findings – System Usability Scale Average

System Usability Scale (SUS)

3. Thinking about your experience interacting with Evidence Library, please respond to the following statements: *

	1 - Strongly Disagree	2	3	4	5 - Strongly Agree
I think that I would like to use this system frequently.	○	○	○	○	○
I found the system unnecessarily complex.	○	○	○	○	○
I thought the system was easy to use.	○	○	○	○	○
I think that I would need the support of a technical person to be able to use this system.	○	○	○	○	○
I found the various functions in this system were well integrated.	○	○	○	○	○
I thought there was too much inconsistency in this system.	○	○	○	○	○
I would imagine that most people would learn to use this system very quickly.	○	○	○	○	○
I found the system very cumbersome to use.	○	○	○	○	○
I felt very confident using the system.	○	○	○	○	○
I needed to learn a lot of things before I could get going with this system.	○	○	○	○	○

The average SUS score for all Evidence Library usability tests for validation is ***80.5****.



*A SUS score above a 68 would be considered ***above average*** and anything below 68 is ***below average***.

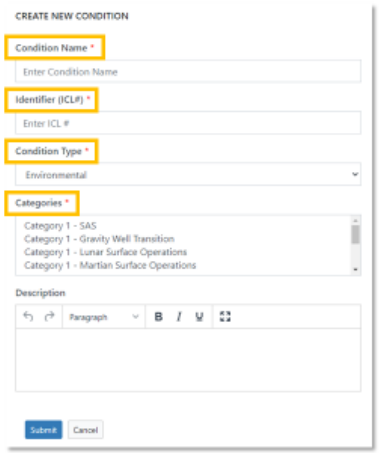
Usability Test Findings and Recommendations

TOC

Scenario E

Red Asterisks

- When creating a new Master Condition, the Name, ICL#, Type, and Category fields all contain red asterisks. However, users can submit/create the condition without entering items in these fields.
- 3 users thought this was inconsistent with the meaning and functionality behind red asterisks, and that the software should provide an error message if items with red asterisks are not entered.
  - 1 of these users mentioned that red asterisks within MedID indicate that the field is required and does not allow the user to save/create items without entering information in the field.
  - 1 of these users indicated that this may cause confusion later on when using search and filter features within Evidence Library.
    - *"If the red asterisks do not work (i.e., they do not provide an error message and restrict users from saving the new condition), it may cause lots of confusion. For example, you wouldn't be able to filter as desired because the tags for that particular condition may not be there."*



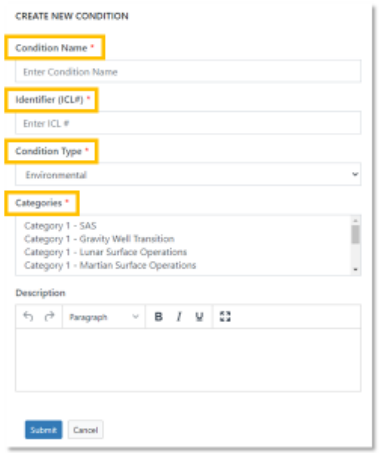
TOC

Scenario E

Red Asterisks

4 Recommendation:

- Red asterisks should be used to indicate required fields.
- If a user does not enter any information, make a selection, etc. for a field with an accompanying red asterisk, provide an error indication/message that informs the user what they need to do before they can submit/create an item.



Usability Test Findings

- **Relevant items for the Transition to Operations Process as defined in Appendix D of NPR 8900.1B:**

> 2. Data demonstrating the efficacy, effectiveness, or utility of the deliverable or product.
> 3. Data demonstrating the operational validation of the deliverable or product.

- **Through an inspection of the data provided in the usability test reports the IMPACT MD team has:**
  - Determined the average SUS score for the Evidence Library user interface, which meets the minimum threshold defined in *PTRS.IMPACT. 099 – System Usability Scale Minimum Score* and therefore demonstrates the usability (i.e., effectiveness, efficiency, satisfaction, and acceptability) is ***acceptable*** and ***excellent***.
  - Reviewed the qualitative data (observations, behaviors, and user feedback) and determined that the product meets users' needs and expectations.
  - Identified the strengths and weaknesses of the IMPACT MD user interfaces and created a plan for implementing HF recommendations for future iterations.

Benefits of Integrating MBSE and HF for V&V Efforts

- **In our HCD efforts to keep our users (active stakeholders) involved and at the forefront throughout the process, MBSE has aided in providing effective usability test evaluations with active stakeholders.**
  - No time was spent in creating tasks/scenarios during planning phase for this software validation activity.
  - Elimination of language barrier that tends to exists between systems engineering language and stakeholders (users). Test participants who had no previous exposure to system modeling languages or the MagicDraw tool were able to effortlessly understand the scenario diagrams.
  - MBSE diagrams allowed test participants to perform each scenario seamlessly and in return, users' feed back was truly focused on the usability of our product and less on understanding the test instructions (task analyses).

Conclusion

- Challenges associated with crew health management for human exploration beyond low Earth orbit continue to grow as we aim to reach new milestones, such as living on Mars.

- As crew members will be unable to rely on real-time communication with Earth-based medical experts due to the lag-time subsequent of the distance between Earth and Mars, addressing complexities rooted in medical systems design and development is vital and takes highly effective multidisciplinary teams to deliver reliable single source of truth data.

- Exploration Medical Capability is a priority that requires advancing medical systems design and risk-informed decision-making for the safety and health of our astronauts.

Conclusion

- **Here, we have shown how SE/MBSE tools (e.g., MagicDraw) and HF methods (e.g., task analyses, usability testing) and approaches (HCD) have been applied to ensure a more optimal end-product.**

- **Using these tools have allowed us to:**
  - Provide a validated and credible single source of truth for medical devices and clinical evidence data used as inputs to the IMPACT tool suite.
  - Communicate and collaborate with other teams across MD more effectively and efficiently.
  - Continue to put our stakeholders at the forefront of our product at every stage in the project.
  - Effectively and timely identify and address gaps that may prevent the user from completing their intended goal.
  - Assess the accuracy of the Scenarios within the ConOps and identify areas for improvement.

Future Work

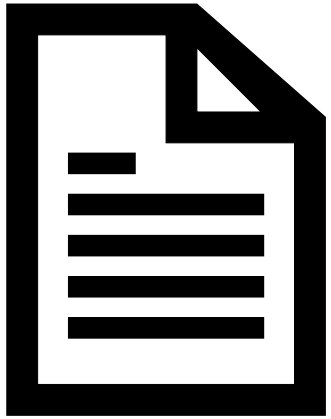
- **Scenario flaws and improvements identified through the V&V efforts of utilizing ConOps scenario will be addressed in upcoming ConOps revision.**
  - As design drivers are strongly dependent on the ConOps, addressing these findings will better capture stakeholder expectations and help facilitate a better understanding of the system.

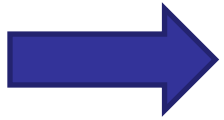
- **Continue to investigate new areas where MBSE approaches and HF methods could be applied.**
  - To aid in providing consistency across products, to allow for more robust documentation, a single source of truth, and effective communication across highly distributed multi-functional teams.
  - To ensure continuous involvement of stakeholders from the beginning to the end of a product's life cycle and understanding their interactions and thoughts. This will aid in ensuring the optimal product is being developed, reduce project costs, and increase user satisfaction.

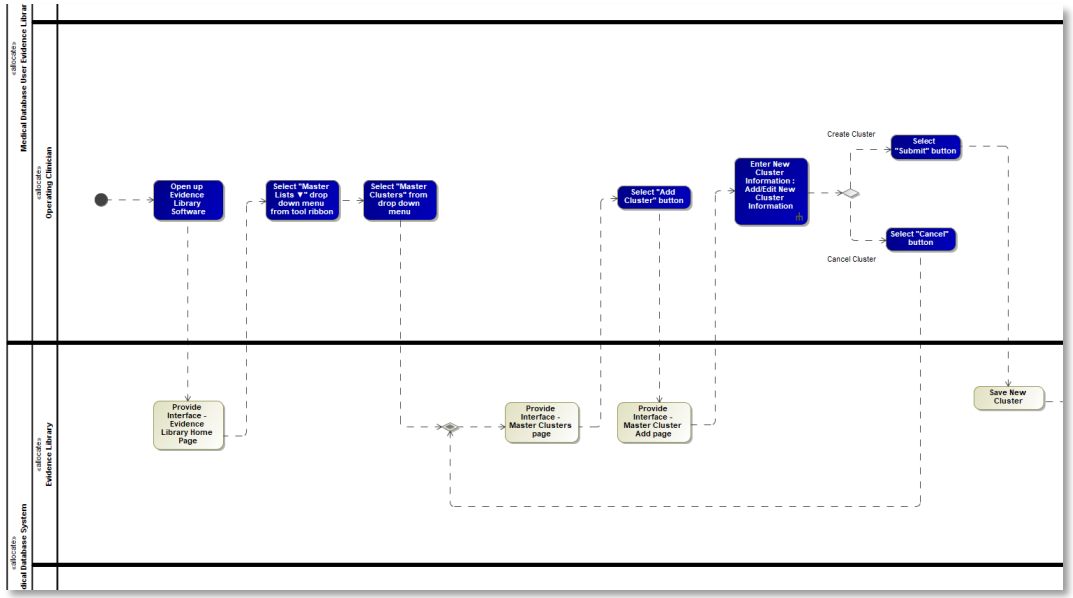

EXMC
EXPLORATION MEDICAL CAPABILITY
NASA
Questions and Comments

What is Model-Based Systems Engineering?

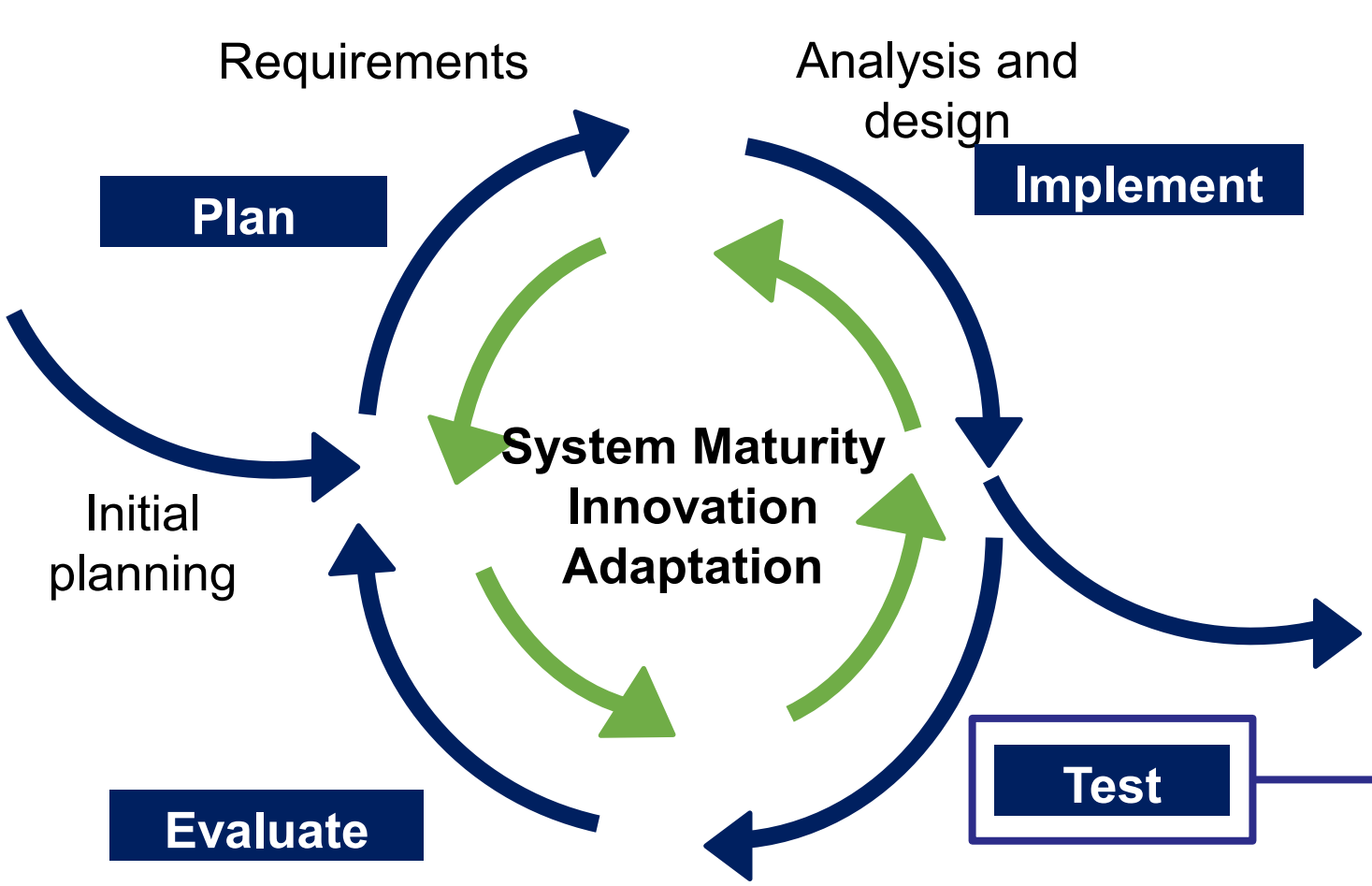
- **Model-based Systems Engineering uses a model of the system to support the systems engineering processes.**
- **Using the model to contain the system information has advantages over traditional document-based approaches.**
- **For example, creating the IMPACT MD Concept of Operations from the system model improved communication with stakeholders.**







What is Human-Centered Design?

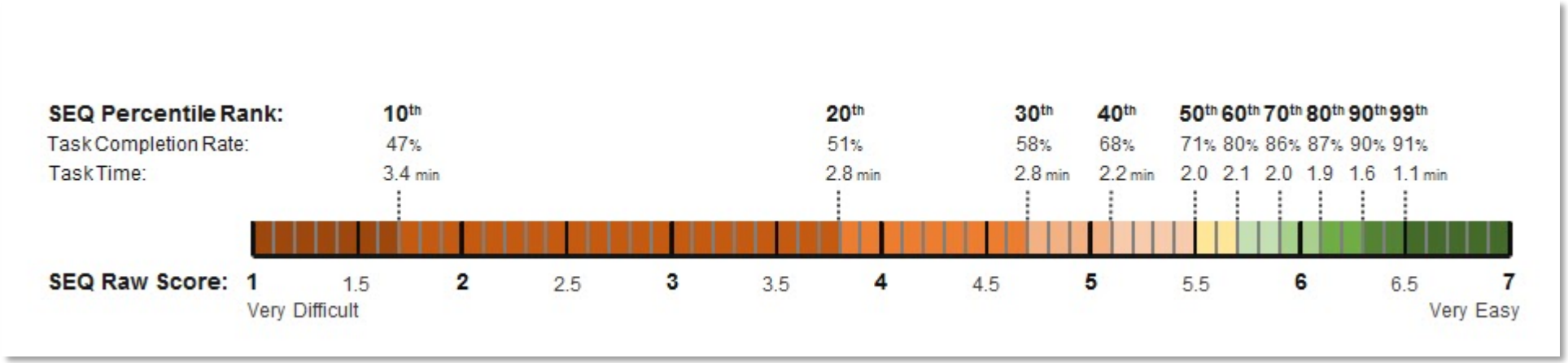


Human Factors Methods

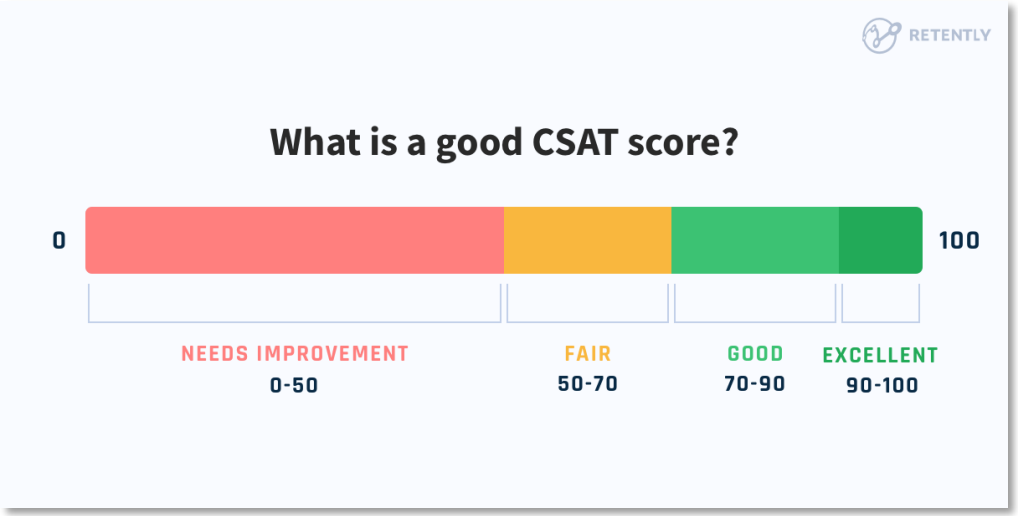
- Task Analyses
- Heuristic Evaluations
- Usability Tests
- A/B Tests
- Interviews
- Cognitive Walkthroughs
- Focus Groups

Ways to Interpret SEQ and CSAT Scores

Single Ease Question (SEQ)



Customer Satisfaction Score (CSAT)



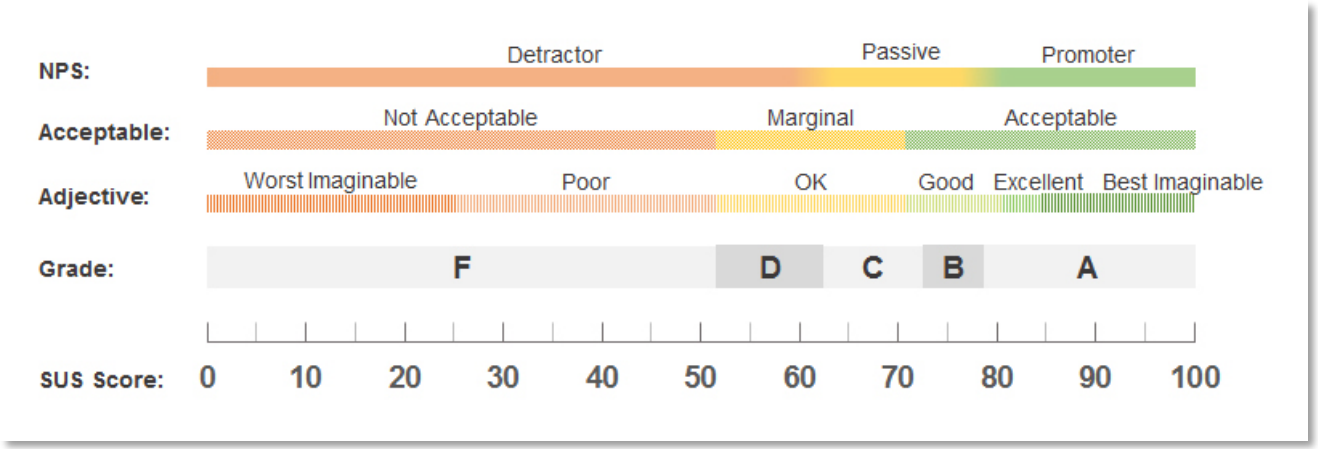
Ways to Interpret System Usability Scale (SUS) Scores

System Usability Scale (SUS)

3. Thinking about your experience interacting with Evidence Library, please respond to the following statements: *

	1 - Strongly Disagree	2	3	4	5 - Strongly Agree
I think that I would like to use this system frequently.	○	○	○	○	○
I found the system unnecessarily complex.	○	○	○	○	○
I thought the system was easy to use.	○	○	○	○	○
I think that I would need the support of a technical person to be able to use this system.	○	○	○	○	○
I found the various functions in this system were well integrated.	○	○	○	○	○
I thought there was too much inconsistency in this system.	○	○	○	○	○
I would imagine that most people would learn to use this system very quickly.	○	○	○	○	○
I found the system very cumbersome to use.	○	○	○	○	○
I felt very confident using the system.	○	○	○	○	○
I needed to learn a lot of things before I could get going with this system.	○	○	○	○	○

A SUS score above a 68 would be considered ***above average*** and anything below 68 is ***below average***.



The average SUS score for all Evidence Library usability tests for validation is ***80.5***.

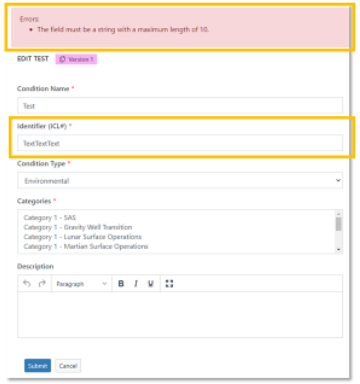
Examples of Findings and Outcomes for Evidence Library

TOC

Scenario E

Error Messages

- When entering a string of text or numbers longer that 10 characters for the *Identifier (ICL#)*, users receive an error message that states:
  - "The field must be a string with a maximum length of 10."
- 1 user did not initially know what field they needed to adjust and indicated that the error message was not helpful in recovering from their error.
  - *"The error message does not indicate what field the issue is originating from, which is not helpful."*



3 Recommendation:

- Ensure that error messages help users recognize, diagnose, and recover from errors.
- Consider editing the text in the error message to state:

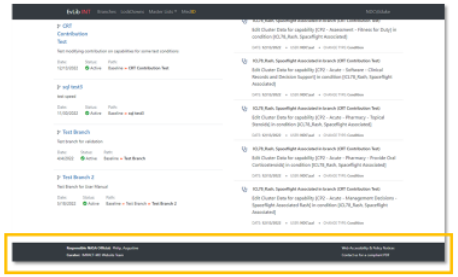
"The Identifier (ICL#) field must be a maximum length of 10 characters."

TOC

Scenario D Scenario E Scenario G

Link for Help Documentation, and Diagrams

- 3 users indicated they would want to see a link somewhere within the Evidence Library user interface where they can access help documentation and the diagrams.
  - *"A help feature that contains the diagrams would be nice to include somewhere within the user interface."*
  - *"I would expect the diagrams to be available in the user manual, which may be accessible from Evidence Library."*



2 Recommendation:

- Incorporate a navigation link that takes users to the user manual and the diagrams within the IMPACT Medical Database model. One suggestion would be to incorporate the link somewhere in the software's bottom tool ribbon.

TOC

Scenario E

Scenario E Description

- 2 users indicated that there is no need for a second clinician to look at condition details at the master level and suggested removing the last sentence of the Scenario E description from the Concept of Operations.
- 1 of these users was not sure why Scenarios E and F are separate.
  - *"Scenario E is just a clinician creating a placeholder and scenario F is actually modifying the CliFF. I would not expect the verifying clinician to validate condition details at the master level as stated in the final sentence of the scenario description."*
  - *"There is no need for a second clinician to look at condition details at the master level."*

Recommendation:

- Needs further discussion.
- No recommendations at this time.

Scenario E – Create New Medical Condition

A user wants to add a new medical condition to the Medical Database that will apply to different studies moving forward. The user opens up the Evidence Library software and navigates to the Master Conditions page. She creates a new Master Condition and inputs all of the information relating to the conditions that applies at the master level. She saves the new master conditions and a second, verifying clinician validates the information that she has input.

TOC

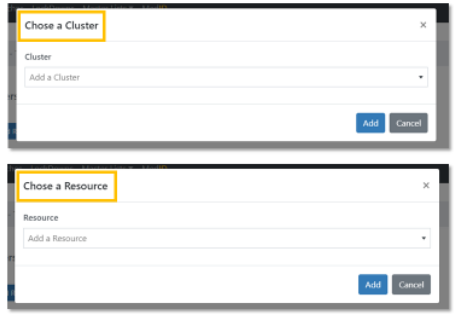
Scenario D

Adding Clusters and Resources

- When adding a cluster or resource to a treatment cluster, 2 users indicated that the spelling on the popup was incorrect. It is spelled "Chose" instead of "Choose".

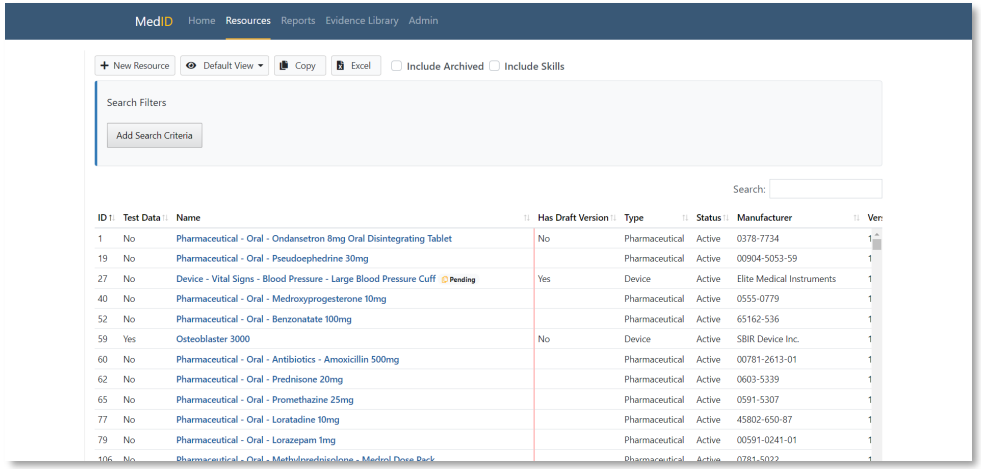
2 Recommendation:

- Correct the spelling in each popup. Change the word from "Chose" to "Choose".



Medical Database Components

MedID



Evidence Library